УДК 004.89
DOI: 10.15827/0236-235X.133.091-097
Дата подачи статьи: 21.12.20
2021. Т. 34. № 1. С. 091–097

Обнаружение аномалий сетевого трафика методом глубокого обучения

В.Н. Зуев [1], *заведующий лабораторией, zvn_tver@mail.ru*

[1] *НИИ «Центрпрограммсистем», г. Тверь, 170024, Россия*

В статье рассматривается применение машинного обучения для обнаружения аномалий в сетевом трафике. В качестве инструмента используются искусственные нейронные сети глубокого обучения. Исследуется эффективность использования нейронных сетей глубокого обучения с использованием набора данных NSL-KDD. Главная особенность этого набора данных – несбалансированность классов.

Описывается метод эффективного использования целевой функции для обучения нейронной сети методом обратного распространения ошибки на несбалансированных примерах. Применение данного алгоритма сопряжено с рядом сложностей, главная из которых – обеспечение приемлемой способности к обобщению нейронной сети. Способность к обобщению полученных знаний является одним из важнейших свойств нейронной сети и заключается в генерации нейронной сетью ожидаемых значений на данных, непосредственно не участвующих в процессе обучения. Однако использование зашумленных и ошибочных данных может привести к переобучению и снижению способности к обобщению обученной нейронной сети.

Предложенный метод позволяет более эффективно рассчитывать значение целевой функции, лежащей в основе алгоритма обратного распространения ошибки. Он хорошо подходит для использования неоднородных выборок при обучении нейронных сетей данных, а также учета при обучении априорной информации о ценности отдельных примеров. В статье приведен алгоритм работы данного метода. Его использование позволяет повысить точность работы нейронной сети для задач классификации и аппроксимации.

Экспериментальные результаты показали, что данный метод хорошо подходит для выявления аномалий в сетевом трафике.

Ключевые слова: *обнаружение вторжений, компьютерная атака, глубокое обучение, аномалии сетевого трафика, обнаружение аномалий, нейронная сеть, машинное обучение, ПК «Ребус-СОВ».*

Как показывает статистика, публикуемая в ежегодном отчете компании CheckPoint, количество известных атак каждый год стремительно растет [1]. Для противостояния данной угрозе необходимо использовать эффективные средства защиты, такие как *системы обнаружения вторжений* (СОВ). Эти средства защиты обычно используют для обнаружения вторжений сигнатурный анализ и требуют регулярного обновления баз сигнатур вторжений. Они не способны обнаруживать атаки, сигнатуры которых отсутствуют в базах сигнатур.

Основанные на обнаружении аномалий методы более перспективны, поскольку могут обнаруживать неизвестные ранее атаки без необходимости предварительного создания сигнатур вторжений для каждой новой атаки. Одно из наиболее актуальных направлений в сфере обнаружения аномалий – выявление аномалий в сетевом трафике.

Любая ЭВМ, функционирующая в локальной вычислительной сети, потенциально подвержена атакам со стороны злоумышленников. По итогам 2018 года средствами ГосСОПКА (государственная система обнаружения, предупреждения и ликвидации последствий компьютерных атак) Российской Федерации было выявлено более 4,3 млрд компьютерных воздействий на критическую информационную инфраструктуру, из них более 17 тысяч наиболее опасных компьютерных атак [2].

Средства обнаружения вторжений, использующие в своей работе только сигнатурный метод обнаружения вторжений, не могут обнаруживать новые виды атак или модификации старых, поэтому актуальна задача разработки алгоритмов выявления аномалий сетевого трафика. Для защиты критически важной инфраструктуры необходимо наличие сертифицированных *Федеральной службой по техническому и экспортному контролю* (ФСТЭК) Российской Федерации средств обнаружения вторжений, способных противостоять современным угрозам.

Одним из таких средств является программный комплекс обнаружения вторжений «Ребус-СОВ» (ПК «Ребус-СОВ») [3], разработанный ЗАО НИИ «Центрпрограммсистем», что

подтверждается сертификатом ФСТЭК России на соответствие документам «Методический документ. Профиль защиты систем обнаружения вторжений уровня сети второго класса защиты» ИТ.СОВ.С2.ПЗ и «Методический документ. Профиль защиты систем обнаружения вторжений уровня узла второго класса защиты» ИТ.СОВ.У2.ПЗ. ПК «Ребус-СОВ» может использоваться на ЭВМ, объединенных в вычислительную сеть и функционирующих под основными ОС, используемыми в ВС РФ, такими как ОС семейства Windows, ОС МСВС и ОС СН «Astra Linux Special Edition».

В данной работе рассмотрена разработка новых методов обнаружения аномалий сетевого трафика с использованием тестового набора данных NSL-KDD.

Описание используемых данных

Существует множество наборов данных для задачи тестирования алгоритмов обнаружения аномалий сетевого трафика. Для исследования был выбран набор NSL-KDD, поскольку он является одним из немногих, использующих протоколы работы TCP, UDP и ICMP, и все его записи разделены на обучающие и тестовые [4]. Набор данных является модернизированной версией набора KDD99, ставшего стандартом для проведения тестов средств обнаружения вторжений. Набор данных NSL-KDD обладает рядом преимуществ по сравнению со стандартным KDD99 [5]:

– удален ряд избыточных и дублирующихся данных с целью устранения их влияния на алгоритмы классификации;

– удалены дублирующиеся записи;

– записи сбалансированно разделены на записи для обучения и записи для тестирования, что исключает необходимость их разделения случайным образом.

Каждая запись в базе NSL-KDD Dataset представляет собой последовательность пакетов, зафиксированную за промежуток времени. Данная последовательность – это поток данных между источником и адресатом сетевых пакетов в соответствии с IP-адресом в заголовке пакета.

Записи включают 41 информационный признак и промаркированы как «атака» и «не атака». База содержит 36 типов атаки, разделенных на четыре категории:

– Denial of Service (dos); набор атак, в которых злоумышленник ограничивает доступ верифицированным пользователям к конкретному сервису через определенный протокол (Back, Land, Neptune, Pod, Smurf, Teardrop, Apache2, Udpstorm, Processtable, Worm);

– Remote to Local (r2l); набор атак, в которых злоумышленник пытается получить доступ извне к локальной машине пользователя (Guess_Password, Ftp_write, Imap, Phf, Multihop, Warezmaster, Warezclient, Spy, Xlock, Xsnoop, Snmpguess, Snmpgetattack, Httptunnel, Sendmail, Named);

– User to Root (u2r); набор атак, в которых злоумышленник, имея доступ к машине жертвы, пытается получить права более привилегированного пользователя (Buffer_overflow, Loadmodule, Rootkit, Perl, Sqlattack, Xterm, Ps);

– Probe; набор атак, в которых злоумышленник пытается получить сведения об инфраструктуре пользователя (Satan, Ipsweep, Nmap, Portsweep, Mscan, Saint).

В NSL-KDD содержатся 125 973 записи для обучения и 22 544 записи для тестирования. Гистограмма распределения представленных в наборе NSL-KDD данных по типам сетевых атак показана на рисунке 1.

Предварительная подготовка данных

Как видно из диаграммы, приведенной на рисунке 1, представленные в наборе данных образы типов атак не сбалансированы. Помимо этого, количество образов трафика, помеченного как «не вторжение», составляет 9 711.

В совокупности с проблемой большой размерности входных и выходных данных неравномерность выборки значительно снижает эффективность методов машинного обучения. Часто используемый для решения данной проблемы метод – исключить из рассмотрения типы сетевых атак, представленные малым количеством образов [6]. Еще одним подходом является использование вместо типов сетевых атак категорий, описанных выше [7]. Такой подход не решает проблему неравномерности выборки, поскольку классы атак тоже распределены неравномерно, но снижает размерность выходных данных с 37 до 5. На практике подобные подходы приведут к неспособности разработанных методов поиска аномалий обнаруживать редкие типы сетевых атак, поэтому в данной работе уменьшение размерности данных рассматривается только за счет уменьшения числа входных факторов.

В работе [8] выполняется проверка значимости входных данных с использованием нечеткой логики. В результате количество факто-

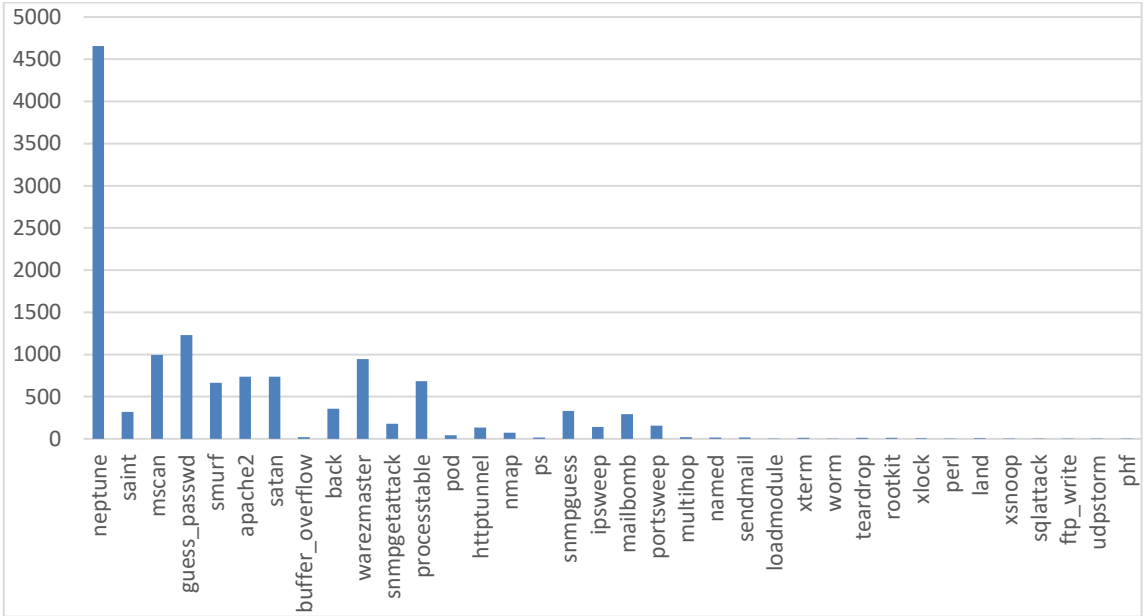

Рис. 1. Гистограмма распределения типов атак, представленных в наборе NSL-KDD

Fig. 1. Histogram of the distribution of attack types represented in the NSL-KDD set

ров сокращается до 8, что повышает быстродействие разработанной модели, но негативно сказывается на точности.

В работе [9] также оценивалось влияние входных факторов, проводилось исследование влияния их количества на точность работы нечеткого классификатора. Авторы выяснили, что уменьшение количества факторов до 30 повышает точность распознавания. Уменьшение количества факторов до 16 приводит к заметному ухудшению точности распознавания.

Наибольшая точность распознавания отмечена в работе [10]. Авторы показали, что атрибуты № 9, 20 и 21 не влияют на точность распознавания, а атрибуты № 7, 8, 11, 14, 15, 17, 19, 32 и 40 оказывают незначительное влияние. При исключении данных факторов получается набор данных с 29 атрибутами.

В описанных работах атаки классифицируются на четыре представленные выше категории. В данной работе исследуется возможность классификации атак по их типам с использованием 29 атрибутов, значимость которых доказана в работе [10].

Перед началом разработки нейросетевой модели обнаружения вторжений необходимо провести предобработку данных. Большинство входных факторов имеют значения 0 и 1, однако некоторые атрибуты, например, время атаки, имеют числовые значения выше единицы. Чтобы устранить влияние таких факторов на процесс обучения нейронной сети, выполняется их нормирование по формуле

$$x' = \frac{x - \min(X)}{\max(X) - \min(X)}, \quad \text{где } x \in X. \qquad (1)$$

Три входных атрибута и выходные значения являются текстовыми. Для использования этих данных произведено кодирование их значений к бинарному виду в соответствии с порядковыми номерами значений. Так, например, значения атрибута protocol type будут закодированы следующим образом: TCP – 0 0, UDP – 0 1, ICMP – 1 1. В таблице 1 приведены количество текстовых значений указанных атрибутов, количество нейронов для их кодирования и значение максимального элемента в бинарном виде.

Таблица 1

Параметры текстовых атрибутов

Table 1

Text Attribute parameters

Параметр	**Количество текстовых значений**	**Количество нейронов для кодирования**	**Бинарное значение максимального элемента**
protocol type	3	2	1 1
service	64	6	1 1 1 1 1 1
flag	11	4	1 0 1 1
Тип атаки	36	6	1 0 0 0 1 1

Используемые методы

Объем данных, обрабатываемый для выявления аномалий, огромен, поэтому для решения данной задачи часто применяются методы машинного обучения, в частности, искусственные нейронные сети.

В последнее время в задаче обнаружения сетевых аномалий активно применяется глубинное обучение – техника обучения нейронных сетей, которая использует множество слоев для решения сложных проблем. Глубокие нейронные сети – это нейронные сети, содержащие несколько скрытых слоев. Такая архитектура доказала свое преимущество при решении различных задач [11]. В работе [12] рассматривается использование глубокого многослойного персептрона, в результате чего достигается средняя точностью 91 %. Авторы предложили распознавать аномалии сетевого трафика следующими глубокими моделями нейронных сетей: рекуррентной нейронной сетью (Recurrent Neural Network, RNN), сложенной рекуррентной нейронной сетью (Stacked RNN) и сверточной нейронной сетью (Convolutional Neural Network, CNN). Для классификации сетевых атак рассмотрены также структуры глубоких нейронных сетей на основе 1D-сверточных и рекуррентных слоев (Long Short-Term Memory, LSTM).

В данной работе рассматривается использование глубокого многослойного персептрона, обучаемого обычным методом и с помощью модифицированного алгоритма обучения, подробное описание которого приведено в [13]. Описанный алгоритм позволяет использовать для обучения нейронных сетей неоднородные выборки данных, а также учитывать при обучении априорную информацию о ценности отдельных примеров.

Суть данного метода заключается в адаптивном присвоении весовых коэффициентов обучающим примерам в зависимости от величины ошибки обучения по данному примеру. Обучение нейронной сети осуществляется методом обратного распространения ошибки, в котором корректировка синаптической карты весов нейронной сети выполняется после подачи всех обучающих примеров, по усредненному значению градиента целевой функции, формулируемой в виде квадратичной суммы разностей между фактическими и ожидаемыми значениями выходных сигналов [14]:

$$E(w) = \frac{1}{2}\sum_{k=1}^{m}(y_k - d_k)^2, \qquad (2)$$

где y – выходное значение нейронной сети; d – желаемое значение выхода; m – количество нейронов в выходном слое; k – номер нейрона в выходном слое.

В пакетном режиме обучения ошибка, рассчитанная по этим примерам, может потеряться в суммарной ошибке по всей выборке, в результате чего такие примеры могут быть проигнорированы.

Для решения данной проблемы обучающим примерам необходимо присвоить весовые коэффициенты. Они будут использоваться при расчете ошибки обучения E и усиливать вклад выбранных примеров в суммарную ошибку обучения.

В начале процесса обучения используются все имеющиеся примеры. Поскольку нейронная сеть перед началом обучения инициализируется случайными значениями, распределение ошибок по используемым примерам равномерное. При этом ошибки по всем примерам не превышают значение E_2. В процессе обучения нейронной сети ошибка обучения приобретает нормальный вид распределения. У большинства примеров ошибка стремится к нулю, у части примеров ошибка остается большой. Попав в интервал $E_1 < E < E_2$, пример получает весовой коэффициент и ошибка по нему начинает уменьшаться быстрее.

Как только ошибка примера пересекает границу E_1, он получает единичный коэффициент. Если при дальнейшем обучении ошибка примера превысит E_1, он снова получит усиливающий весовой коэффициент. Если ошибка примера превысила значение E_2, то пример считается выбросом и не участвует в дальнейшем обучении. Направления изменений ошибок показаны пунктирными стрелками на рисунке 2.

Граничные значения E_1, E_2 и весовые коэффициенты рассчитываются индивидуально для каждого примера на каждом шаге обучения.

Эксперименты и оценка результатов

Согласно результатам проведенных исследований, структура сети для данной задачи представляет собой однонаправленную нейронную сеть с двумя скрытыми слоями, состоящую из нейронов сигмоидального типа (многослойный персептрон). Передача сигнала в этой сети осуществляется только в одном направлении от входа к выходу.

Входной слой содержит 38 нейронов. Первый и второй скрытые слои содержат по 20 нейронов, а в выходном слое находятся 6 ней-

ронов. Структура нейронной сети для выявления аномалий сетевого трафика приведена на рисунке 3.

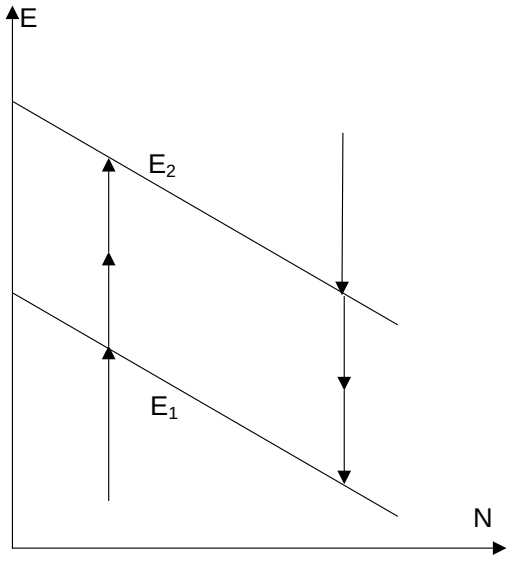

Рис. 2. Направления изменений ошибок выхода в процессе обучения

Fig. 2. Directions of changes in output errors in the learning process

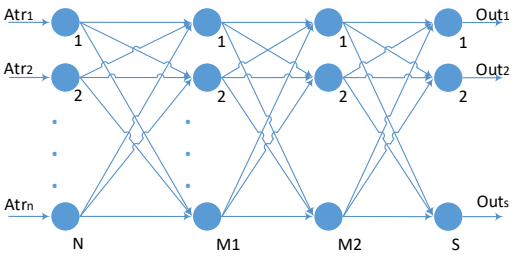

Рис. 3. Структура нейронной сети для выявления аномалий сетевого трафика

Fig. 3. The neural network structure for detecting network traffic anomalies

Проверка разработанных моделей на тестовой выборке показала среднюю точность классификации сетевых вторжений 91 % для модели, обучаемой стандартным методом, и 89 % для модели, обучаемой с помощью модифицированного алгоритма. Если же рассмотреть отдельно типы вторжений, представленных малым количеством образов, то точность модели, обученной с помощью модифицированного алгоритма, выше более чем в два раза. Точность классификации вторжений двух рассмотренных моделей для типов, представленных малым количеством образов, приведена в таблице 2.

Таблица 2

Точность классификации вторжений, представленных малым количеством образов

Table 2

Accuracy of classification of intrusions represented by a small number of images

Вторжение	Качество образа	Точность классификации стандартным методом, %	Точность классификации модифицированным алгоритмом, %
sendmail	14	40	71
loadmodule	2	50	100
xterm	13	46	77
worm	2	0	100
teardrop	12	60	83
rootkit	13	30	84
xlock	9	55	89
perl	2	50	100
land	7	42	86
xsnoop	4	50	75
sqlattack	2	0	100
ftp_write	3	33	67
udpstorm	2	50	100
phf	2	0	100

Заключение

Разработанная нейросетевая модель выявления сетевых аномалий полностью не исключает возможность ложных срабатываний и пропусков вторжений, однако показала эффективность при обнаружении типов вторжений, представленных малым количеством образов. Метод продолжает развиваться и проходит апробацию в ПК «Ребус-СОВ».

В данной работе предложен новый подход к построению модели выявления аномалий в сетевом трафике, основанный на применении нейронных сетей. Проведенные исследования свидетельствуют о его эффективности. Использование в СОВ данного метода совместно с сигнатурным анализом событий повысит эффективность обнаружения вторжений, в том числе принципиально новых и модифицированных существующих.

Для повышения степени эффективности обнаружения вторжений необходимо продолжение исследования. Основной задачей дальнейшей работы является уменьшение количества ложных срабатываний, которые все еще могут возникать при функционировании на реальных объектах.

Литература

1. Check Point Security Report 2019. URL: https://blog.checkpoint.com/2019/03/04/check-points-2019-security-report// (дата обращения: 07.11.2020).
2. Национальный координационный центр по компьютерным инцидентам. URL: http://cert.gov.ru/ (дата обращения: 07.11.2020).
3. Программный комплекс обнаружения вторжений «Ребус-СОВ». URL: https://rebus-sov.ru/ (дата обращения: 07.11.2020).
4. NSL-KDD Dataset. URL: https://www.unb.ca/cic/datasets/nsl.html (дата обращения: 07.11.2020).
5. Bhattacharjee P., Fujail A., Begum S. Intrusion detection system for NSL-KDD data set using vectorised fitness function in genetic algorithm. Advances in Computational Sciences and Technology, 2017, vol. 10, pp. 235–246.
6. Guojie L., Jianbiao Z. Research of network intrusion detection based on convolutional neural network. Discrete Dynamics in Nature and Society, 2020, vol. 2020, pp. 1–11. DOI: 10.1155/2020/4705982.
7. Dhanabal L., Shantharajah S.P. A study on NSL-KDD dataset for intrusion detection system based on classification algorithms. IJARCCE, 2015, vol. 4, no. 6, pp. 446–452.
8. Choudhary S., Kesswani N. Analysis of KDD-Cup'99, NSL-KDD and UNSW-NB15 datasets using deep learning in IoT. Procedia Computer Science, 2020, vol. 167, pp. 1561–1573. DOI: 10.1016/j.procs.2020.03.367.
9. Paulauskas N., Auskalnis J. Analysis of data preprocessing influence on intrusion detection using NSL-KDD dataset. Proc. Open Conf. of eStream, 2017, pp. 1–5. DOI: 10.1109/eStream.017.7950325.
10. Ingre B., Yadav A. Performance analysis of NSL-KDD dataset using ANN. Proc. Intern. Conf. SPACES, 2015, pp. 92–96. DOI: 10.1109/SPACES.2015.7058223.
11. Гафаров Ф.М., Галимянов А.Ф. Искусственные нейронные сети и приложения. Казань: Изд-во Казан. ун-та, 2018, 121 с.
12. Chockwanich N., Visoottiviseth V. Intrusion Detection by deep learning with TensorFlow. Proc. XXI ICACT, 2019, pp. 654–659. DOI: 10.23919/ICACT.2019.8701969.
13. Зуев В., Кемайкин В. Модифицированный алгоритм обучения нейронных сетей // Программные продукты и системы. 2019. Т. 32. № 2. С. 258–262. DOI: 10.15827/0236-235X.126.258-262.
14. Оссовский С. Нейронные сети для обработки информации; [пер. с польск.]. М.: Финансы и статистика, 2002. 344 с.

Software & Systems
DOI: 10.15827/0236-235X.133.091-097

Received 21.12.20
2021, vol. 34, no. 1, pp. 091–097

Network anomalies detection by deep learning

V.N. Zuev [1], *Head of Laboratory, zvn_tver@mail.ru*

[1] *R&D Institute Centerprogramsystem, Tver, 170024, Russian Federation*

Abstract. The paper discusses the machine learning application for detecting anomalies in network traffic. Artificial neural networks of deep learning are used as a tool. In this paper, the NSL-KDD data set is analyzed and used to study the effectiveness of deep learning neural networks in detecting anomalies in network traffic patterns. The most important aspects of this dataset are the imbalanced class distribution.

The paper describes the method of effective usage of objective functions backpropagation algorithms in order to train the neural network on imbalanced samples. Using the backpropagation algorithm is connected with many difficulties. The major problem is the ability to generalize the neural network. The ability to generalize is the most important characteristic of a neural network. It is mean that trained on studying data neural network is capable to produce output value by using unknown data. However, using for training noisy data decreases the ability to generalize the neural network.

The proposed method makes it possible to more efficiently calculate the value of the aim function, which is the basis of the error back-propagation algorithm. The method is well fit for the heterogeneous sample and can use priority information about the sample's significance. The pepper described an algorithm of the method. Using this method will improve the accuracy of the neural network for classification and regression problems.

The experimental result shows that it well suits the designed method for network anomaly detections.

Keywords: intrusion detection, computer attack, deep learning, network traffic anomalies, anomaly detection, neural network, machine learning, PC "Rebus-SOV".

References

1. *Check Point Security Report 2019*. Available at: https://blog.checkpoint.com/2019/03/04/check-points-2019-security-report// (accessed November 07, 2020).

2. *The National Coordination Center for Computer Incidents*. Available at: http://cert.gov.ru/ (accessed November 07, 2020).

3. *Rebus-SOV Intrusion Detection Software Package*. Available at: https://rebus-sov.ru/ (accessed November 07, 2020).

4. *NSL-KDD Dataset*. Available at: https://www.unb.ca/cic/datasets/nsl.html (accessed November 07, 2020).

5. Bhattacharjee P., Fujail A., Begum S. Intrusion detection system for NSL-KDD data set using vectorised fitness function in genetic algorithm. *Advances in Computational Sciences and Technology*, 2017, vol. 10, pp. 235–246.

6. Guojie L., Jianbiao Z. Research of network intrusion detection based on convolutional neural network. *Discrete Dynamics in Nature and Society*, 2020, vol. 2020, pp. 1–11. DOI: 10.1155/2020/4705982.

7. Dhanabal L., Shantharajah S.P. A study on NSL-KDD dataset for intrusion detection system based on classification algorithms. *IJARCCE*, 2015, vol. 4, no. 6, pp. 446–452.

8. Choudhary S., Kesswani N. Analysis of KDD-Cup'99, NSL-KDD and UNSW-NB15 datasets using deep learning in IoT. *Procedia Computer Science*, 2020, vol. 167, pp. 1561–1573. DOI: 10.1016/ j.procs. 2020.03.367.

9. Paulauskas N., Auskalnis J. Analysis of data preprocessing influence on intrusion detection using NSL-KDD dataset. *Proc. Open Conf. of eStream*, 2017, pp. 1–5. DOI: 10.1109/eStream.017.7950325.

10. Ingre B., Yadav A. Performance analysis of NSL-KDD dataset using ANN. *Proc. Intern. Conf. SPACES*, 2015, pp. 92–96. DOI: 10.1109/SPACES.2015.7058223.

11. Gafarov F., Galimjanov A.F. *Artificial Neural Networks and Applications.* Kazan, 2018, 121 p. (in Russ.).

12. Chockwanich N., Visoottiviseth V. Intrusion Detection by deep learning with TensorFlow. *Proc. XXI ICACT*, 2019, pp. 654–659. DOI: 10.23919/ICACT.2019.8701969.

13. Zuev V., Kemajkin V. An improved neural network training algorithm. *Software and Systems*, 2019, vol. 32, no. 2, pp. 258–262 (in Russ.). DOI: 10.15827/0236-235X.126.258-262.

14. Ossovskij S. *Neural Networks for Information Processing*. Moscow, 2002, 344 p. (in Russ.).

Для цитирования

Зуев В.Н. Обнаружение аномалий сетевого трафика методом глубокого обучения // Программные продукты и системы. 2021. Т. 34. № 1. С. 091–097. DOI: 10.15827/0236-235X.133.091-097.

For citation

Zuev V.N. Network anomalies detection by deep learning. *Software & Systems*, 2021, vol. 34, no. 1, pp. 091–097 (in Russ.). DOI: 10.15827/0236-235X.133.091-097.